

## 遺伝的アルゴリズムのナップザック問題への 適用における特徴の解析

謝 孟春\* 山口 哲司\* 小高 知宏\* 小倉 久和\*

### An Analysis of Characteristics of the Genetic Algorithm in Application to the Knapsack Problem

Mengchun XIE, Tetsushi YAMAGUCHI, Tomohiro ODAKA, and Hisakazu OGURA

(Received Feb. 28, 1994)

Genetic Algorithm(GA) is an innovative search algorithm based on natural phenomena evolution, whose main advantages lie in great robustness and problem independence. Among all the algorithms used in the field of optimization problems, GA is quite successful one. In this paper we study the genetic operators, i.e., crossover, mutation, selection, etc., for the knapsack problem. In order to yield good results, we make use of the informations of the nonliving genes which are produced in the genetic operations by designing the fitness functions for nonliving genes. We discussed the relationship between the results obtained and the fitness function for nonliving genes. In addition, to examine the evolutionary states we introduce an expression for the variety of chromosomes in the population, and propose a heuristic mutation operator based on the analysis of the relation between the strict solution and the ratio value/weight.

## 1 はじめに

遺伝的アルゴリズム (Genetic Algorithm, GA) は、1970 年代に Holland[1] が生物進化における遺伝子の役割を模倣して提唱したものである。その枠組みが、適用分野に比較的依存しない、汎用的なアルゴリズムであるため、近年その応用研究が盛んに行われるようになってきた。GA がランダムな確率的探索と区別されるのは、操作論的には交叉や突然変異といった遺伝的操作と、理論的には染色体の優秀な断片の結合と成長が進化の主役であるとするビルディングブロック仮説である。しかし、GA の数学的基礎は未だほとんど確立していないため、遺伝子操作がどのような特徴をもち、進化においてどのような役割を果たすのか、また、ビルディングブロックそのものをどのように表現するか、その形成と成長が世代とともにどのように進化していくのか、よくわかっていないのが現状である。

ナップザック問題とは、異なる重量と価値を併せ持つ複数の荷物が与えられたとき、重量の合計が制限重量を越えない範囲で価値の合計が最大になるようにいくつかの荷物の選び方を求める、という制約条件付き組合せ最適化問題である。この極めて単純な組み合わせ最適化問題であるナップザック問題は、GA の応用が困難な問題とされている。ナップザック問題におけるビルディングブロックが染色体の一部を占める訳ではなく、一般にはモザイク的に形成されるため、遺伝的操作がその特徴を

\*工学部情報工学科

発揮しにくいのではないかと考えられる。本論文では、このナップザック問題を対象に、GA の特徴を検討する。

交叉、突然変異、選択・淘汰という基本的な遺伝的操作方法はいろいろ提案されているが、どのような方法がナップザック問題によりよく適用するのかは自明ではない。第2章では、ナップザック問題に GA を適用するためのモデルを設定し、このような遺伝的操作についていくつかの方法を試み、その効果を検討する。

制約条件を満足しない解候補に対応する染色体は致死遺伝子となる。しかし、致死遺伝子にもビルディングブロックは形成されており、致死遺伝子が進化過程に参加することにより優秀な個体が生成される可能性がある。第3章では、非致死遺伝子の島（此岸）と致死遺伝子の島（彼岸）からなる此岸-彼岸モデルを提案し、致死遺伝子の進化における影響を考察する。

GA では進化につれて集団の多様性が変化する。多様性は局所解からの脱出の保証であるが、高速な収束の妨げでもある。第4章では、この集団の多様性を表現する方法を提案し、検討する。また、この表現方法により示されるナップザック問題の一般的特徴を突然変異の生起に関するヒューリスティックとして取り入れる試みについて述べる。

## 2 ナップザック問題と遺伝的操作

### 2.1 GA のナップザック問題への適用

ナップザック問題において、荷物の数を  $N$  とすると、荷物の選び方は  $2^N$  通りであり、その中で制約条件を満たす選び方だけが解となる。解の中で価値の合計が最大なものを厳密解という。理論的には  $NP$  完全問題であり、一般に厳密解を多項式時間で求めるアルゴリズムは知られていない。

ナップザック問題は、荷物の総数を  $N$  として、 $i$  番目の荷物の重さと価値をそれぞれ  $w_i$ 、 $v_i$ 、制限重量を  $LIMIT$  として、

$$V = \sum_{i=1}^N v_i x_i$$

に対して

$$W = \sum_{i=1}^N w_i x_i \leq LIMIT$$

の条件下で  $V$  を最大とする  $(x_1, x_2, \dots, x_N)$  の組合せを決定する問題と定式化できる。ただし、 $x_i$  は、 $i$  番目の荷物の状態を表し、 $x_i$  が 1 の場合は荷物を詰め、 $x_i$  が 0 の場合は荷物を詰めないことを意味する。

$$x_i \in \{0, 1\}, \quad i = 1, 2, \dots, N$$

$N$  個の荷物を対象とするナップザック問題における一つの解候補を一つの染色体で表わす。染色体は遺伝子の列からなり、遺伝子座（遺伝子の染色体での位置）を荷物の番号に対応させる。ナップザックにその荷物を詰めるかどうかで、対応する遺伝子座に遺伝子 0 または 1 を割り振る。したがって、遺伝子型としての染色体は長さが  $N$  の 0 と 1 の並びであるビット列となる。例えば、荷物の数が全部で 10 個であり、1、3、8、9 番目の荷物が選択された解候補は、遺伝子型として 1 0 1 0 0 0 0 1 1 0、その表現型は 1 3 8 9 となる。

長さ  $N$  の任意のビット列が解候補になり得るが、重量の合計が制限重量を越える個体は、問題の有効な解に対応しない致死遺伝子となる。

$N$  個の荷物を扱うナップザック問題を考える。まず、個体の集団（初期集団）を構成する。1 つの個体は、各遺伝子の値をランダムに 0 か 1 に選んで長さ  $N$  の染色体を形成することによって生成される。このような個体の集団を遺伝子プールという。初期集団は  $K$  個の個体からなる。集団中の各

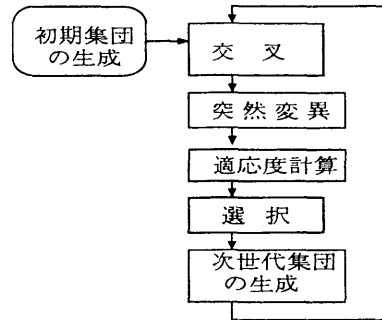


図 1: 遺伝的操作の流れ

染色体は、ランダムに相手を選んで交叉を行い、子供を集団に残すことにより、増殖を行う。次に、集団からいくつかの個体をランダムに選び、突然変異を施す。どのような交叉と突然変異の方法を用いるかということは GA にとって本質的な課題であり、次節で考察する。

こうして初期集団から増殖した集団は、次に選択・淘汰される。ある個体  $k$  の適応度  $f_k$  ( $k = 1, 2, \dots, K$ ) をその個体の遺伝子から求められる価値の総和、

$$f_k = \begin{cases} \sum_{i=1}^N v_i x_i, & (W \leq LIMIT) \\ 0, & (W > LIMIT) \end{cases}$$

とする。制限重量を越えた致死遺伝子に対しては、適応度は 0 とする。各個体は適応度に応じて生き残り選択され、残りは淘汰されて、遺伝子プールのサイズは初期集団と同じサイズ  $K$  となる。さらに、本研究では、エリート保存戦略も用いている。これは、遺伝子操作前の集団中で最大の適応度を持つ個体を無条件に生き残り選択する方法である。選択の戦略も GA にとって大きな影響がある。選択法についても次節で検討する。

ある個体集団から始めて、交叉、突然変異、選択・淘汰を行って、再び同じサイズの個体集団を得る。これを世代交代という。初期集団を第 0 世代として、このような世代交代を次々に行うことにより、第  $P$  世代 ( $P = 1, 2, \dots$ ) が得られる。

## 2.2 遺伝的操作の考察

遺伝的操作とは古い個体から新しい個体を作る方法である。ナップザック問題に対して、交叉、突然変異、選択などの基本的な遺伝的操作を繰り返すことによって、最適解 (または近似最適解) を得ようとする。一般に 2 つの個体の遺伝子のコードが類似しているからといってその表現型あるいは適応度が類似しているとは限らない。遺伝子のコードが変化しても表現型や適応度が殆ど変化しない場合もあるし、逆に極めて大きな変化を引き起こす場合もある。したがって、一般に遺伝的操作の方法は、GA を適用した結果に大きな影響を与える。この節では、ナップザック問題に GA を適用した試行結果に基づき、選択、交叉、突然変異の遺伝的操作について、考察し、評価する。

### (1) 選択法

ここでは、2つの選択戦略、ルーレット選択法と順序選択法を比較検討する。

ルーレット選択法は各個体をその適応度に比例した確率で選択し、 $K$ 個生き残らせる方法である。この方法では、個体の適応度が高いほど選択されるチャンスが多いが、適応度が低い個体も選ばれる可能性がある。順序選択法では、集団中の全ての個体を適応度の値の順に並べ、上位の個体  $K$  個を生き残らせる。この選択方法は淘汰圧力が強いので、適応度の低い個体は確実に淘汰され、適応度の高い個体だけが残ることになる。そうすると、集団の遺伝子の組成変化に強い指向性をもたらし、収束が早いと思われるが、反面、集団の多様性が失われてしまい局所的な解に陥る可能性もある。

この二つの選択方法を用いて、荷物の数  $N = 25, 100$  について 200 世代で終了として、解を求める過程における最大適応度の遷移を図 2 に示す。図は 10 回の試行の平均で、破線はルーレット選択法、実線は順序選択法の結果である。図からわかるように、順序選択法はルーレット選択法に比べて、集団中の全体個体の適応度がより高く、適応度の高い個体に応じている問題の解を良く見つけることができる。

しかし、順序選択法では、適応度の高い個体が急速に増え、早い時期に集団の多様性が失われてしまうことがある。何世代か経過した後、個体集団の遺伝子が殆ど同一になってしまうのである。特に集団のサイズが小さいほど、このようなことが早い世代で起る。この場合の例の探索過程を図 3 に示す。左図は荷物が 25、集団のサイズが 25 の例であるが、26 世代で進化が止まった。右図は荷物 50 で、集団のサイズも 50 の場合であるが、39 世代で進化が止まった。その時、集団の染色体は一種類になってしまっている。もっともこのような状態は多かれ少なかれどのような手法でも生じていると考えられるが、検討を要する課題である。

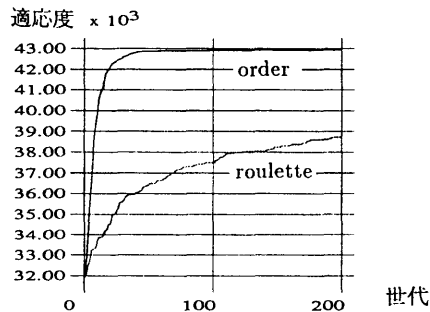


図 2: 二つの選択方法の適応度の遷移

### (2) 交叉法

一点交叉と一様交叉の2つの交叉方法を比較検討する。交叉の対象となる染色体対は常に異なる対となるように選択する。

一点交叉では、二つの染色体をランダムに選んだ位置で交叉させ、遺伝子を入れ換えて二つの子供をつくる。一様交叉では、あらかじめ染色体と同じ長さのランダムな 0,1 の系列からなるマスクを用意しておき、マスクが 1 の遺伝子座の遺伝子を、二つの染色体の間で入れ替えて、子供をつくる。

一点交叉と一様交叉では、染色体を組み換えるパターン数が大きく異なる。である。一点交叉では、交叉点の数の  $N - 1$  通りの組み換えパターンがあるのに対して、一様交叉では、 $2^{N-1}$  通りのマスクのパターンの組み換えが可能で、生成される探索領域が異なる。ナップザック問題に対して、染色体によって遺伝される情報は、荷物の詰め方である。一点交叉では、交叉前後で、交叉点の片側の

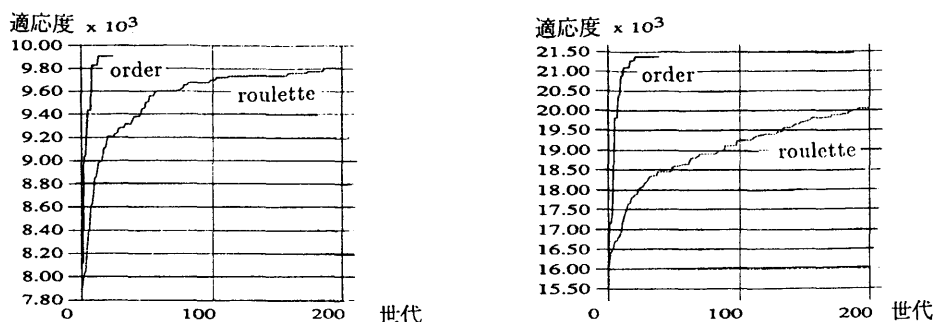


図 3: 順序選択法における進化の停滞

染色体の遺伝子の組合せは変化しない。一点交叉では、探索空間でみると、現在の探索点に近い所を探索点とすることになる。多様なビルディングブロックを形成させようとする立場から見ると、このような方法は適当ではない。

ビルディングブロックは個体の適応度を高めるのに貢献する部分的な形質である。ナップザック問題におけるビルディングブロックは厳密解と一致する染色体中の遺伝子の並びである。一様交叉はランダムに生成されたマスクの各ビットの値によって、親の遺伝子の子供に継承させる方法を決定するから、次の世代に、多様な子孫を生成できる。つまり、集団の多様性を維持できると期待される。

二つの交叉方法を用いて、それぞれの荷物の数に対して、200 世代まで、100 回ずつの試行を行った。1 回の試行で、最終的いくつかの解を得ることができるが、その中で最も優秀解をその試行で得た解とする。よって、100 回の試行により 100 個の解が得られる。その 100 個の解中で、最も良い解 (best1~best5) と最も悪い解 (worst1~1worst5) を表 1 に示す。表の結果を見ると、荷物数  $N = 25$ 、100 の時、一様交叉からより良い解が得られた ( $N = 25$  の時、一点交叉により得た厳密解は 10 回、一様交叉により得た厳密解は 12 回である)。他の場合ではあまり大きな差はみられない。

表 1: 二つの交叉法の結果

荷物数	一点交叉				一様交叉			
	25	50	100	150	25	50	100	150
厳密解	9923	21381	42993	66138	9923	21318	42993	66138
best1	9923	21214	40605	60280	9923	21025	40638	59273
best2	9923	20939	40440	59979	9923	21017	40187	59106
best3	9923	20910	40212	59927	9923	20925	39910	59075
best4	9923	20853	40134	59832	9923	20899	39854	58922
best5	9923	20851	40131	59663	9923	20866	39778	58991
worst1	9448	19252	37153	54411	9364	18650	37318	55435
worst2	9466	19292	37481	54578	9368	19468	37322	55517
worst3	9515	19464	37500	54582	9428	19542	37419	55574
worst4	9522	19592	37600	54638	9440	19546	37518	55767
worst5	9535	19599	37610	55182	9469	19575	37610	55929

### (3) 突然変異法

反転変異と異点交換変異の二つの突然変異方法を比較検討する。

反転変異は、変異を発生する個体と変異点をランダムに指定し、変異点で個体の遺伝子に相当するビットを、0 を 1、あるいは 1 を 0 に変更する方法である。異点交換変異は選んだ個体に対して、二つの変異点を選択し、それらで遺伝子を交換する方法である。

反転変異と異点交換変異の二つの方法では、突然変異による個体の重量と価値の変化の仕方に違いがある。反転は、荷物を新たに追加するか、あらかじめ詰められている荷物を取り出すかという操作であるため、突然変異の対象となる荷物の重量と価値が、そのまま突然変異前と突然変異後の変化量となる。これに対して異点交換は、詰められた荷物を取り出し、別の荷物を新たに追加するという操作なので、突然変異前と突然変異後の変化量は、突然変異の対象となる二つの荷物の差となる。つまり、平均的に見ると、反転は異点交換に比べて個体の重量と価値の変化量が大きくなる。また、反転では重量と価値は同時に増減するのに対して、異点交換は対象となる荷物の重量と価値によっては、重量は減少し価値が増加する（または、その逆）といったことが起こり得る。

ところで、何世代か経過し、集団がある程度成長した段階では、集団中の各個体の重量は制限重量  $LIMIT$  の近くに分布するようになる。致死遺伝子でない個体に対して反転を行う場合、遺伝子の値が 1 から 0 に反転すると、突然変異の対象となる荷物の価値だけ個体の価値の合計が減少し、個体の適応度が下がる。逆に遺伝子の値を 0 から 1 に反転すると、価値の合計は増加するが、それにとともに重量も必ず増加する。このため重量が  $LIMIT$  に近い個体では、重量が  $LIMIT$  を越えて致死遺伝子となる可能性が高い。これに対して、異点交換では、反転に比べて重量の平均的な増加量が少ないので、致死遺伝子になる可能性が低い。しかし、重量と同様に価値の平均的な増加量も反転より少ないので、個体の適応度が飛躍的に高くなることは期待できない。

二つの突然変異の方法を用いて、200 世代まで、それぞれの荷物に対して、10 回ずつ実行中に最良解、最悪解、および 10 回の解の平均値を表に示す。表の結果からわかるように、荷物数が 100 の場合を除いて、反転変異が異点交換変異に比べてより良い解が得られた。これは反転変異を用いると、重量が減少し価値が増加する染色体がよく生成されるためだと考えられる。

表 2: 二つの突然変異法の結果

荷物数	反転変異				異点交換変異			
	50	100	150	200	50	100	150	200
best	20991	40076	59981	75652	20607	40187	58816	74757
worst	19588	38233	55003	71176	19742	37630	56099	70499
average	20291	39132	58095	72900	20224	38722	57161	72715

## 3 致死遺伝子を用いた此岸—彼岸モデル GA

前章で述べた交叉と突然変異という遺伝的操作を行う際に、制限重量  $LIMIT$  を越えた致死遺伝子を発生する可能性がある。通常の GA ではこれを淘汰するが、しかし、一概に致死遺伝子といっても、遺伝子型データに着目すれば、なかには優秀な遺伝子の生成に貢献し得る部分を有するものが存在するはずである。すなわち、致死遺伝子は厳密解の染色体と一致する部分的な構造を持っている可能性がある。交叉と突然変異を行う時、そのようなビルディングブロックが保存され、他の部分が変化されると、非致死遺伝子になって、優秀な個体が生成される可能性が十分にある。本章では、致死遺伝子を淘汰せず、積極的に用いた GA を検討する。

### 3.1 致死遺伝子を用いる GA のモデル

ここで提案する致死遺伝子を用いた GA のモデルでは、まず非致死遺伝子が住んでいる“此岸”と致死遺伝子が住んでいる“彼岸”という二つの島を作る。“此岸”は非致死遺伝子だけに存在を許す。“彼岸”では致死遺伝子だけが生存できる。処理の流れは図4のようになる。

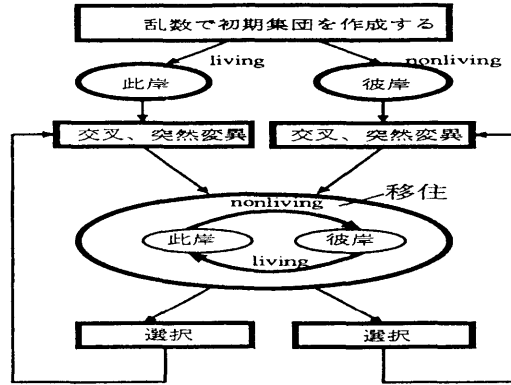


図4: 致死遺伝子を用いた GA の流れ

此岸—彼岸モデルでは、前章で述べた一様交叉と異点交換変異を用いて、“此岸”と“彼岸”で別々に交叉や突然変異などの遺伝的操作を行なう。交叉と突然変異を行った後に、二つの島の設定した環境に相応しない個体が出てくる場合がある。すなわち、“此岸”で致死遺伝子が発生し、“彼岸”で遺伝子が復活して非致死遺伝子となる場合がある。個体の重量をチェックし、“此岸”で発生した致死遺伝子を“彼岸”に移住させ、同じように、“彼岸”で復活した個体を“此岸”に移住させる。移住は適応度と無関係に個体の重量だけで決める。個体は各々の適応する環境に着いてから、評価値に比例した確率で選択される。このようにして、形質の異なる遺伝子を持つ個体が拡散・増殖することができ、遺伝子の情報を十分に利用することができると考えられる。

評価関数は優秀なビルディングブロックを含む個体により高い適応度が与えられるように設定する必要がある。“此岸”での評価関数は2.1節と同じように設定すればよいが、“彼岸”では、どのような致死遺伝子が優秀なのかという判断は容易ではない。そこで、次の四つの方法を比較する。

- 方法1：直線型の重み付けによる評価法  $f1 = \frac{w_m - w_e}{w_m - w_l} c_e$
- 方法2：自乗型の重み付けによる評価法  $f2 = \frac{(w_m - w_e)^2}{w_m - w_l} c_e$
- 方法3：LIMITとの差の逆数による評価法  $f3 = \frac{1}{w_e - w_l} c_e$
- 方法4：“0”評価法  $f4 = \frac{w_m - w_e}{w_m - w_l} c_d$

ここで、 $w_e = \sum_{i=1}^n w_i x_i$ 、 $c_e = \sum_{i=1}^n c_i x_i$ 、 $c_d = \sum_{i=1}^n c_i (1 - x_i)$  である。 $w_m$ は遺伝子の最大重量、 $w_f$ は制限重量 LIMIT、 $x_i$ と  $c_i$  は前の定義と同じである。

方法1では、遺伝子の価値の和だけを用いるのではなく、価値と重量との関係も重視し、なるべく制限重量に近いところの遺伝子の評価値が高くなるようにする。方法2では、重量の重みが大きくなるようにする。方法3は、重量の少ない個体ほど高い適応度を与えようというものである。致死遺伝子である個体は染色体の中に、遺伝子を“1”とする数が多い。方法4では、遺伝子を“0”とする個体により高い適応度を与える。

### 3.2 此岸—彼岸モデル GA の試行結果

致死遺伝子を用いた GA と、致死遺伝子を全て淘汰した GA によって探索した結果を表 3 に比較して示す。試行結果は、それぞれの問題に対して 100 回の試行をおこなったもので、それぞれの試行は 200 世代までおこなう。試行の対象とするナップサック問題には、荷物数が 25、50、100、150 の 4 種類を用意し、荷物の重量と価値の設定は、次のようになっている。

- 各荷物の重量は 5～200、価値は 100～1000 の整数とする。
- *LIMIT* の値は全荷物の総重量の半分とする。
- 各島の遺伝子プールのサイズは荷物数と同じとする。

表 3 から、致死遺伝子をうまく利用することにより解の精度を上げることが可能であることが推測できる。

表 3: 異なる遺伝的操作の比較

設定	荷物の数	25	50	100	150
	厳密解	9923	21381	42993	66138
保留法	厳密解の数	2	0	0	0
	<i>best</i>	9923	21047	40177	59441
	<i>average</i>	9553	19372	37025	54760
淘汰法	厳密解の数	1	0	0	0
	<i>best</i>	9923	20669	39493	58632
	<i>average</i>	9425	19155	36707	54602

“彼岸”で設定した四つの評価関数のどれが致死遺伝子を評価するのに適当かを確認するため実験した結果を表 4 に示す。パラメータの設定は、荷物数が 150 個、各島の初期集団のサイズが 150、他の設定は上と同じである。なお、厳密解は 66138 である。

この表から、探索結果が“彼岸”の評価関数に強く依存することがわかる。評価関数として、自乗の重みを付ける方法 2 は、他の方法に比べよりよい解が得られている。“彼岸”では、異なる評価関数によって異なる形質を持っている致死遺伝子を一時保留することになる。これらの致死遺伝子は最適解に対して異なった方法で貢献をしていると思われる。

表 4: 各試行における最適値の分布

	最高値	誤差率	平均値	最低値
方法 1	59378	0.102	57045	54946
方法 2	59441	0.101	57258	55702
方法 3	58074	0.122	55213	51565
方法 4	55117	0.167	52472	50550



## 4 集団の多様性の検討

### 4.1 集団多様性の表現

これまで試行によって得られた表面的な結果について述べてきたが、以下では個体の内部的表現である染色体 (遺伝子型データ) に着目し、世代が経過するにつれて、染色体がどのように変化していくかを、集団の多様性と関連させて観察する。集団中の個体の染色体の様子を、例えば、次のような記号列を用いて表現することにする。

10\*\*0i01\*\*1o\*\*io\*1

この記号列は染色体と同じ長さであり、各記号は染色体の同じ位置にあたる遺伝子座と対応している。集団中の全個体に対して、各遺伝子座の遺伝子の値を調べた結果によって、その遺伝子座に対応する記号を決める。記号の意味は次のようにである。

- 1 … 各個体の遺伝子のうち 90% 個以上が ‘1’ であり、厳密解と一致している
- 0 … 各個体の遺伝子のうち 90% 個以上が ‘0’ であり、厳密解と一致している
- i … 各個体の遺伝子のうち 90% 個以上が ‘0’ であり、厳密解と一致していない
- o … 各個体の遺伝子のうち 90% 個以上が ‘0’ であり、厳密解と一致していない
- \* … それ以外

このようにして得られる記号列は、集団の様子を表現している。例えば記号列中に ‘\*’ が多く含まれていれば、集団の多様性が維持されていることがわかる。また、‘\*’ が少なく ‘i’ や ‘o’ が多く含まれていれば、局所解へ収束していることになる。

この記号列を用いて、荷物数が 25、200 の場合に、2 章で述べた遺伝子操作のうち一様交叉、異点交換変異、ルーレット選択法を行った場合の各世代の集団プールにおける染色体の様子を表すと次のようになる。

#### 《荷物数 25 の場合》

```

4 世代 *1*****1*****
6 世代 *1*****1***0*****
8 世代 *1*****10***1***0*0****
10 世代 *1***11**1****1*0*0*0****
13 世代 111*111**1*****0*0*0o*
15 世代 *11***1**1**0*110***01***
18 世代 1110**11*1**0*11**0*010*1
20 世代 11101111*1****11**0*010*1
23 世代 111*1111*10*0*11**0*010o1
25 世代 111*1111*10*0*110*01010*1
30 世代 111*1111*10*0*110*01010*1
35 世代 111*1111*10*0*110101010*1
40 世代 11101111010***110*01010*1
45 世代 111*11110*0*0*110101010*1
50 世代 111*1111010*0*11010*010*1

```

#### 《荷物数 200 の場合》

```

5 世代 *****0*****
10 世代 *****1***o***1*****0*****1*****0*0*****
15 世代 *****1***1***o***1*****101**1*****1*****0*0***0..

```

```

20 世代 ****11*1****i**11**oo01*o****1*0*****1*011*****1****1**0*0o****
25 世代 0*1111***0*1***111i*o01**o***1*0**i1*1*011i*****1**0*1*10*0o****
30 世代 011111*1*0o1i1*111ioo01**o*****0**i1*1*011i*****1*0o1110*0oo*0
40 世代 *11111i**0o1i1*111io*0101o1*1*0**i01*011i11*1*00*1*10*0o****
50 世代 011111i1*0*1i1*111*o*0101o1*10011i10**011i11***00*1110*0o**0
60 世代 011111i110*1i1*111*o*0101o1*10011*1010011i11*1*00*1110o0o*00
70 世代 011111i11011i1*111***0101o1*1*011*1010011i11*11*0*1110o0o*00
80 世代 011111*11011i1*1110*o010*o1*1*01101010011i11*11*011110o0**00
90 世代 011111011011i*i1110*o010**101i01101010011i11011*011110o0***0
100 世代 011111011011i1i11101*0101*101i01101010011i11011*011110o01100
150 世代 011111011011i1i11101*0101o101i01101010011i110110011110o01*00
200 世代 011111*11011*1*1110110101*101*01101010011i110110011110o01100
250 世代 0111110110110101110110101o101i01101010011i110110011110o01*00

```

このように初期の世代は‘\*’の記号が多いが、世代を経過するごとに‘\*’の数は急速に減少する。このことから、突然変異によって集団中の個体の種類数は保たれているが、それぞれの個体の染色体は世代が経過するにつれて非常に似かよってくるのがわかる。荷物数 25 の場合は‘i’や‘o’が、ほとんど現れないが、荷物数が多くなるにしたがって‘i’や‘o’が多く現れるようになる。特に初期の段階において多く現れる。図 5 は各世代の‘\*’の数(点線)と‘i’と‘o’の合計数(実線)をグラフにしたものである。左図は荷物数が 25 であり、右図は荷物数が 200 である。

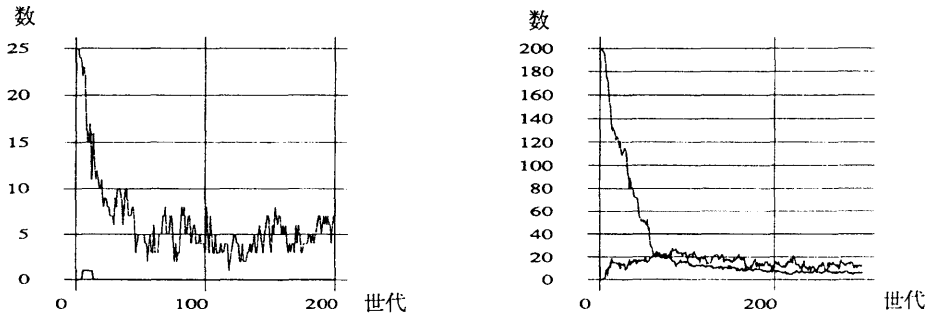


図 5: 集団の多様性

このように‘\*’の数は荷物数に関係なく 50 世代あたりまでに急速に減少し、集団の多様性が失われる。このとき荷物数 25 の場合のように‘i’や‘o’がなければ、うまく厳密解を求めることができる。荷物数 200 の場合のように多様性が失われた時点で‘i’や‘o’がたくさんあると、以後の世代でその数を減らすことはできても最終的にはいくつか残り、局所的な収束から脱却するのは困難である。‘i’や‘o’があると必ずしも厳密解を得ることができないわけではないが、‘i’や‘o’があるということは、対応する遺伝子座の遺伝子が厳密解と異なる値に収束していることになるので、このような集団の状態からでは厳密解を求めるのは困難である。

なぜこのように荷物数が多くなると初期の段階で‘i’や‘o’が多く現れるようになるのかを検討する。図 6 は、荷物数 200 の試行におけるエリート個体の染色体と厳密解を表す染色体とのハミング距離を表すグラフである。縦軸がそのハミング距離であり、横軸は正の方向に向かうほど遅い世代

に現れたエリート個体に対応している。この図からわかるように初期の段階におけるエリート個体の染色体は厳密解を表す染色体とのハミング距離が大きく、あとに現れるほどハミング距離は小さくなる。このため初期の段階で現れるエリート個体の形質が選択、交叉によって集団中に広まると、厳密解と異なる遺伝子が集団中で支配的になり、この結果‘i’や‘o’がたくさん現れる。あとに現れるエリート個体ほど厳密解とのハミング距離が小さいので、その形質が集団中に広まった結果‘i’や‘o’の数が減少する。しかし、進化が停滞し始めると新たなエリート個体が現れなくなり、その時点でのエリート個体の形質が集団中で強く支配的になり集団の多様性が失われてしまう。このような状態になると交叉が有効に作用しなくなり、それ以上の進化には突然変異がうまく具合に起こるなどの偶然が必要となる。

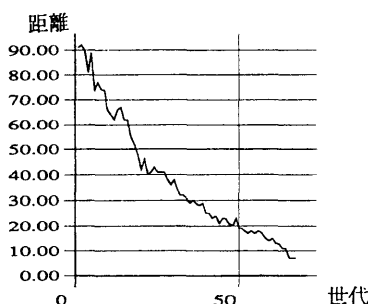


図 6: 厳密解との hamming 距離

## 4.2 ナップザック問題の特徴を利用した突然変異操作

GA の実際問題への応用において、問題の特別な知識、または問題の性質を用いなくても、最適な解が得られることは GA の大きな特徴である。しかし、問題の知識と特性を GA に取り入れれば、それを利用することでより探索能力が向上する GA を構築することができる可能性がある。そこで、前節の結果から得られたナップザック問題の知識を取り入れた GA について考察する。

厳密解を表す染色体の遺伝子座を、対応する荷物の (価値/重量) の値の大きい順に並べ変えると染色体は次ようになる。

荷物数 50    11111...11111101111100000000...00000

このように荷物数に関係なくどの染色体も左端から 1、右端から 0 が並び、1 と 0 の境のかなり狭い範囲で 1 と 0 が入り混じって並んでいる。そして、前節で説明した集団中の染色体の様子を表す記号列を染色体と同じように遺伝子座を並べ変え、上の厳密解を表す染色体と照らし合わせると次のようになる。

荷物数 50    ...1111111111110001000000...    厳密解  
                   ...1\*111111\*1\*000\*0\*00\*0...    500 世代目の記号列

次に、荷物数 50 の場合の並べ換えを行なった記号列を世代順に並べると次のようになる。

5 世代    \*1\*\*\*\*\*0\*\*\*\*\*  
 10 世代    11\*\*\*1\*\*\*1\*\*1\*\*\*\*\*0\*\*\*\*\*0\*\*\*\*\*

```

15 世代 111*11*1*1*11*****1**o*****0****0*0**00*0*0*0
20 世代 111111111111111111**1*****1*1*****0****0*0**00*000*0
25 世代 111111111111*111*11****1*1o*****0****0*0*000i0*0*0
30 世代 111111111111*11*1****1*1*****0****0*0*0*0*00000
35 世代 1111111111111111*1***111*o*1o0i*o0****0*000*00000
40 世代 111111*11111*111*11*11111o**o0i*o0****0*000*0*00*
45 世代 11111111111111111111**1**11o01o0i*o0***0*0*0*0*0000*
50 世代 111111111111*111111*11**11o0i*0***o0*00*00000000000
55 世代 111111111111*111111*111*11o0i*0i**0**0000000*00000
60 世代 111111*11111*111111*11*11*o0i*0i**0*0000000*000000

```

このように染色体の両端から‘\*’以外の記号が現れはじめ、世代が経過するにしたがって徐々に染色体の中央部にも‘\*’以外の記号が現れてくる。つまり、染色体の両端に位置する遺伝子座の遺伝子ほど初期の世代で収束する傾向があることになる。しかも、これらの遺伝子が厳密解と異なる値に収束することはほとんどなく、収束してもその状態がそのままずっと続くこともほとんどない。このようなナップサック問題の性質は一般的なもので、多くの場合はこのような性質をもっていると期待される。もちろん、全く逆の状況となるも考えられる。

上のようなナップサック問題の性質を突然変異に対するヒューリスティックとして、次のようにしてGAに取り入れることを考える。まず対応する荷物の(価値/重量)の値の大きい順に遺伝子座を並べ変えたものを個体の染色体とする。このような染色体をもつ個体集団に対して、今までと同じように選択、交叉をおこなう。交叉で生成された  $K$  個の新しい個体集団に対して、‘1’、‘0’、‘\*’の3つの記号列を上と同様に生成する。それぞれの記号の意味は次のようである。

- 1 …  $K$  個の個体の遺伝子がすべて ‘1’ である
- 0 …  $K$  個の個体の遺伝子がすべて ‘0’ である
- \* … それ以外

このような条件で生成した記号列を用いて、突然変異を起こす遺伝子座の限定をおこなう。生成した記号列に、左端から ‘1’ が連続している部分、あるいは右端から ‘0’ が連続している部分があれば、その連続している部分に対応する遺伝子座を突然変異を起こす対象から除く。例えば、ある世代における遺伝子プールに対して次のような記号列が生成されたとする、

```
111**1.....0*000*
```

この集団における突然変異は4番目以降の遺伝子座を対象とすることになる。以上の手続きを各世代で突然変異を操作する前におこなう。したがって、突然変異の対象となる遺伝子座は、世代を経過するごとに徐々に減ってゆくことになる。

このようなヒューリスティックを利用した突然変異と単純な突然変異操作の比較を表5に示す。ただし致死遺伝子をもつ個体は淘汰し、交叉は一様交叉、突然変異は異点交換を採用する。打ち切り世代は500世代、試行回数は50である。どの荷物数においても、よい結果を得ることができた。また、厳密解が求まらなかった試行でもより厳密解に近い解を得ている。荷物数200で厳密解を得た回数が他の場合よりもかなり少なくなっているが、これは、厳密解の染色体において1と0が入り混じっている範囲が広いことが原因だと考えられる。

表 5: 試行結果

荷物数	ヒューリスティックを利用する				ヒューリスティックを利用しない			
	50	100	150	200	50	100	150	200
厳密解	1697	4006	6278	8474	1697	4006	6278	8474
worst 1	1695	4006	6268	8465	1695	3994	6255	8438
worst 2	1695	4006	6271	8465	1695	3995	6255	8443
worst 3	1697	4006	6271	8465	1695	3995	6257	8445
worst 4	1697	4006	6271	8466	1695	3995	6258	8445
worst 5	1697	4006	6273	8466	1697	3995	6260	8446
平均	1696.92	4006.00	6276.86	8470.10	1696.24	4001.40	6268.16	8459.02
厳密解数	48	50	38	5	31	16	9	0

## 5 考察

本論文では、ナップザック問題に対する GA の適用において、いくつかの検討と提案を行った。まず、ナップザック問題における遺伝的操作の方法について検討した。また、此岸-彼岸モデルを提案し、致死遺伝子におけるビルディングブロックを積極的に利用する方法についても検討した。さらに、GA における 1 つの重要な指標である多様性に関する表現方法を提案し、それを利用したナップザック問題のヒューリスティックを利用した GA を試みた。

遺伝的操作は、遺伝子コーディングの方法に大きく依存する。本研究で採用した二進コーディングにおいては、問題の性質からみれば、交叉では一様交叉、突然変異では異点交換が基本的には有利であると考えられ、ここで報告しなかった試行結果もそのような傾向を示しているが、しかし、それは統計的なレベルであり、決定的に有利であるということにはならなかった。これがどのようなことを意味するかは不明であるが、他のコーディング方法でどうなるかを含めて、今後の検討課題として残されている。

通常、致死遺伝子は無条件に淘汰される。一般に、致死遺伝子は、ある個体の染色体が、遺伝子コーディングの手法に反映できていない制約条件を満足しないために発生するのであるから、染色体上のどこか一部あるいは全部が「致死」を引き起こしているのである。実際、ある致死遺伝子の遺伝子コーディングを最適解の遺伝子コーディングと比較すれば部分的にはビルディングブロックを含んでいることがある。ここで扱ったナップザック問題の設定では、ほとんどの致死遺伝子は有効なビルディングブロックを含んでいる。第 3 章で提案・検討した此岸-彼岸モデルの試行結果は、致死遺伝子を積極的に扱うことにより、このように「死に体」となったビルディングブロックの利用が可能となることを示していると思われる。しかし、致死遺伝子をどのように評価することがよいのかは自明でない。ここでは制限重量との差の自乗に逆比例する評価関数を中心にいくつかの手法を比較検討したが、これは今後の重要な課題の一つである。

GA のシミュレーションにおいて、集団の染色体の様子をどのように表現するかは容易ではない。たとえば、初期集団ではさまざまな形をもつ染色体があるけれども、進化につれて集団中の染色体の種類が少なくなってくるが、そのようすをどう表現するのかということである。第 4 章で述べた手法によれば、それがどのように進んでいくのか簡単に手に取るようになる。さらに工夫することによりビルディングブロックがどのように成長していくか観察する手段となると考えられる。しかし、この手法のこのような特徴はナップザック問題を二進コーディングしたための特徴であり、GA でいつもこのような表現が可能なのではない。一般の GA におけるビルディングブロックの成長過程を

観察・測定する手段を検討する必要がある。

4.2 節で示した簡単なモデルシミュレーションの結果によれば、突然変異の対象となる遺伝子座に関するヒューリスティックによって突然変異確率を操作することは、探索結果に大きなもたらす。このような突然変異確率の操作を工夫することにより、さらに改善が期待できる。

## 参考文献

- [1] Holland, J.H.: "Adaptation in Natural and Artificial Systems", The Univ. Michigan Press(1975), MIT Press(1992)
- [2] Syswerda, G.: "Uniform Crossover in Genetic Algorithms", Proc. of ICGA-89, pp2-9(1989)
- [3] Grefenstette John J.: "Optimization of Control Parameters for Genetic Algorithms", IEEE Trans. Systems, Man, and Cybernetics, Vol. SMC-16 No.1, pp122-128(1986)
- [4] Goldberg D.E. and Rudnick M.: "Genetic Algorithms and the Variance of Fitness", Complex Systems, Vol.5 No.3, pp265-278(1991)
- [5] 北野 宏明: 遺伝的アルゴリズム, 産業図書 (1993)
- [6] 和田 健之介: "進化システム論 遺伝的アルゴリズムの基礎 (1)～(7)", Computer Today, No.47～No.54, (1992.1～1993.3)
- [7] 小林 重信: "遺伝的アルゴリズムの現状と課題", 計測と制御, Vol.32 No.1, pp.2-9(1993.1)
- [8] 波田野 寿昭: "遺伝的アルゴリズム", 人工知能学会誌, Vol.8 No.3, pp312-319(1993.5)
- [9] 謝 孟春、小高 知宏、小倉 久和: "ナップザック問題への適用における GA の課題", 平成5年度電気関係学会北陸支部連合大会論文集, p.357(1993.9)